SYSTEM AND METHOD FOR AUTOMATIC MULTIPLE LISTENER ROOM ACOUSTIC CORRECTION WITH LOW FILTER ORDERS

BY

SUNIL BHARITKAR

CHRIS KYRIAKAKIS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The contents of this application are continuation in part of the application filed June 20, 2003 and related to provisional application having serial number 60/390,122 (filed June 21, 2002).

[0002] BACKGROUND

[0003] 1. Field of the Invention

[0004] The present invention relates to multi-channel audio and particularly to the delivery of high quality and distortion-free multi-channel audio in an enclosure.

[0005] 2. Description of the Background Art

[0006] The inventors have recognized that the acoustics of an enclosure (e.g., room, automobile interior, movie theaters, etc.) play a major role in introducing distortions in the audio signal perceived by listeners.

[0007] A typical room is an acoustic enclosure that can be modeled as a linear system whose behavior at a particular listening position is characterized by an impulse response, $h(n)$ $\{n=0, 1, \ldots, N-1\}$. This is called the room impulse response and has an

associated frequency response, $H(e^{j\omega})$. Generally, $H(e^{j\omega})$ is also referred to as the room transfer function (RTF). The impulse response yields a complete description of the changes a sound signal undergoes when it travels from a source to a receiver (microphone/listener). The signal at the receiver contains consists of direct path components, discrete reflections that arrive a few milliseconds after the direct sound, as well as a reverberant field component.

[0008] It is well established that room responses change with source and receiver locations in a room. A room response can be uniquely defined for a set of spatial co-ordinates (x_i, y_i, z_i). This assumes that the source (loudspeaker) is at origin (0, 0, 0) and the receiver (microphone or listener) is at the spatial co-ordinates, x_i , y_i and z_i, relative to a source in the room.

[0009] Now, when sound is transmitted in a room from a source to a specific receiver, the frequency response of the audio signal is distorted at the receiving position mainly due to interactions with room boundaries and the buildup of standing waves at low frequencies.

[0010] One mechanism to minimize these distortions is to introduce an equalizing filter that is an inverse (or approximate inverse) of the room impulse response for a given source-receiver position. This equalizing filter is applied to the audio signal before it is transmitted by the loudspeaker source. Thus, if $h_{eq}(n)$ is the equalizing filter for $h(n)$, then, for perfect equalization $h_{eq}(n) \otimes h(n) = \delta(n)$; where $\otimes$ is the convolution operator and $\delta(n)$ is the Kronecker delta function.

[0011] However, the inventors have realized that at least two problems arise when using this approach, (i) the room response is not necessarily invertible (i.e., it is not minimum phase), and (ii) designing an equalizing filter for a specific receiver (or listener) will produce poor equalization performance at other locations in the room. In other words, multiple-listener equalization cannot be achieved with a single equalizing filter. Thus, room equalization, which has traditionally been approached as a classic inverse filter problem, will not work in practical environments where multiple-listeners are present.

[0012] Furthermore, it is required that for real-time digital signal processing, low filter orders are required. Given this, there is a need to develop a system and a method for correcting distortions introduced by the room, simultaneously, at multiple-listener positions using low filter orders.

[0013] SUMMARY OF THE INVENTION

[0014] The present invention provides a system and a method for delivering substantially distortion-free audio, simultaneously, to multiple listeners in any environment (e.g., free-field, home-theater, movie-theater, automobile interiors, airports, rooms, etc.). This is achieved by means of a filter that automatically corrects the room acoustical characteristics at multiple-listener positions.

[0015] Accordingly, in one embodiment, the method for correcting room acoustics at multiple-listener positions comprises: (i) measuring a room acoustical response at each listener position in a multiple-listener environment; (ii) determining a general response by computing a weighted average of the room acoustical responses; and (iii) obtaining a

room acoustic correction filter from the general response, wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions. The method may further include the step of generating a stimulus signal (e.g., a logarithmic chirp signal, a broadband noise signal, a maximum length signal, or a white noise signal) from at least one loudspeaker for measuring the room acoustical response at each of the listener position.

[0016] In one aspect of the invention, the general response is determined by a pattern recognition method such as a hard c-means clustering method, a fuzzy c-means clustering method, any well known adaptive learning method (e.g., neural-nets, recursive least squares, etc.), or any combination thereof.

[0017] The method may further include the step of determining a minimum-phase signal and an all-pass signal from the general response. Accordingly, in one aspect of the invention, the room acoustic correction filter could be the inverse of the minimum-phase signal. In another aspect, the room acoustic correction filter could be the convolution of the inverse minimum-phase signal and a matched filter that is derived from the all-pass signal.

[0018] Thus, filtering each of the room acoustical responses with the room acoustical correction filter will provide a substantially flat magnitude response in the frequency domain, and a signal substantially resembling an impulse function in the time domain at each of the listener positions.

[0019] In another embodiment of the present invention, the method for generating substantially distortion-free audio at multiple-listeners in an environment comprises: (i)

measuring the acoustical characteristics of the environment at each expected listener position in the multiple-listener environment; (ii) determining a room acoustical correction filter from the acoustical characteristics at the each of the expected listener positions; (iii) filtering an audio signal with the room acoustical correction filter; and (iv) transmitting the filtered audio from at least one loudspeaker, wherein the audio signal received at said each expected listener position is substantially free of distortions.

[0020] The method may further include the step of determining a general response, from the measured acoustical characteristics at each of the expected listener positions, by a pattern recognition method (e.g., hard c-means clustering method, fuzzy c-means clustering method, a suitable adaptive learning method, or any combination thereof). Additionally, the method could include the step of determining a minimum-phase signal and an all-pass signal from the general response.

[0021] In one aspect of the invention, the room acoustical correction filter could be the inverse of the minimum-phase signal, and in another aspect of the invention, the filter could be obtained by filtering the minimum-phase signal with a matched filter (the matched filter being obtained from the all-pass signal).

[0022] In one aspect of the invention, the pattern recognition method is a c-means clustering method that generates at least one cluster centroid. Then, the method may further include the step of forming the general response from the at least one cluster centroid.

[0023] Thus, filtering each of the acoustical characteristics with the room acoustical correction filter will provide a substantially flat magnitude response in the frequency

domain, and a signal substantially resembling an impulse function in the time domain at each of the expected listener positions.

[0024] In one embodiment of the present invention, a system for generating substantially distortion-free audio at multiple-listeners in an environment comprises: (i) a multiple-listener room acoustic correction filter implemented in the semiconductor device, the room acoustic correction filter formed from a weighted average of room acoustical responses, and wherein each of the room acoustical responses is measured at an expected listener position, wherein an audio signal filtered by said room acoustic correction filter is received substantially distortion-free at each of the expected listener positions. Additionally, at least one of the stimulus signal and the filtered audio signal are transmitted from at least one loudspeaker.

[0025] In one aspect of the invention, the weighted average is determined by a pattern recognition system (e.g., hard c-means clustering system, a fuzzy c-means clustering system, an adaptive learning system, or any combination thereof). The system may further include a means for determining a minimum-phase signal and an all-pass signal from the weighted average.

[0026] Accordingly, the correction filter could be either the inverse of the minimum-phase signal or a filtered version of the minimum-phase signal (obtained by filtering the minimum-phase signal with a matched filter, the matched filter being obtained from the all-pass signal of the weighted average).

[0027] In one aspect of the invention, the pattern recognition means may be a c-means clustering system that generates at least one cluster centroid. Then, the system

may further include means for forming the weighted average from the at least one cluster centroid.

[0028] Thus, filtering each of the acoustical responses with the room acoustical correction filter will provide a substantially flat magnitude response in the frequency domain, and a signal substantially resembling an impulse function in the time domain at each of the expected listener positions.

[0029] In another embodiment of the present invention, the method for correcting room acoustics at multiple-listener positions comprises: (i) clustering each room acoustical response into at least one cluster, wherein each cluster includes a centroid; (ii) forming a general response from the at least one centroid; and (iii) determining a room acoustic correction filter from the general response, wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions.

[0030] In one aspect of the present invention, the method may further include the step of determining a stable inverse of the general response, the stable inverse being included in the room acoustic correction filter.

[0031] Thus, filtering each of the acoustical responses with the room acoustical correction filter will provide a substantially flat magnitude response in the frequency domain, and a signal substantially resembling an impulse function in the time domain at the multiple-listener positions.

[0032] In another embodiment of the present invention, the method for correcting room acoustics at multiple-listener positions comprises: (i) clustering a direct path component

of each acoustical response into at least one direct path cluster, wherein each direct path cluster includes a direct path centroid; (ii) clustering reflection components of each of the acoustical response into at least one reflection path cluster, wherein said each reflection path cluster includes a reflection path centroid; (iii) forming a general direct path response from the at least one direct path centroid and a general reflection path response from the at least one reflection path centroid; and (iv) determining a room acoustic correction filter from the general direct path response and the general reflection path response, wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions.

[0033] In another embodiment of the present invention, the method for correcting room acoustics at multiple-listener positions comprises: (i) determining a general response by computing a weighted average of room acoustical responses, wherein each room acoustical response corresponds to a sound propagation characteristics from a loudspeaker to a listener position; and (ii) obtaining a room acoustic correction filter from the general response, wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions.

[0034] In another embodiment of the present invention, the method for correcting room acoustics at multiple-listener positions using low order room acoustical correction filters comprises the steps of: (i) measuring a room acoustical response at each listener position in a multiple-listener environment; (ii) warping each of the room acoustical response measured at said each listener position; (iii) determining a general response by computing a weighted average of the warped room acoustical responses; (iv) generating a low order spectral model of the general response; (v) obtaining a warped

acoustic correction filter from the low order spectral model; and (vi) unwarping the warped acoustic correction filter to obtain a room acoustic correction filter; wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions. The method may further including the step of generating and transmitting a stimulus signal (e.g., an MLS sequence, a logarithmic-chirp signal) for measuring the room acoustical response at each of the listener positions. The general response could be determined by a weighted average approach (as in through a pattern recognition method). The pattern recognition method could at least one of a hard c-means clustering method, a fuzzy c-means clustering method, or an adaptive learning method. The warping may be achieved by means of a bilinear conformal map. The spectral model includes at least one of a pole-zero model and Linear Predictive Coding (LPC) model. The warped acoustic correction filter is the inverse of the low order spectral model.

[0035] In another embodiment, a method for generating substantially distortion-free audio at multiple-listeners in an environment comprises: (i) measuring acoustical characteristics of the environment at each expected listener position in the multiple-listener environment; (ii) warping each of the acoustical characteristics measured at said each expected listener position; (iii) generating a low order spectral model of each of the warped acoustical characteristics; (iv) obtaining a warped acoustic correction filter from the low order spectral model; (v) unwarping the warped acoustic correction filter to obtain a room acoustic correction filter; (vi) filtering an audio signal with the room acoustical correction filter; and (vii) transmitting the filtered audio from at least one

loudspeaker, wherein the audio signal received at said each expected listener position is substantially free of distortions.

[0036] The system for generating substantially distortion-free audio at multiple-listeners in an environment comprises: a filtering means for performing multiple-listener room acoustic correction, the filtering means formed from: (a) warped room acoustical responses, wherein the room acoustical responses are measured at each of an expected listener position in a multiple-listener environment; (b) a weighted average response of the warped room acoustical responses; (c) a low order spectral model of the weighted average response; (d) a warped filter formed from the low order spectral model; and (e) an unwarped room acoustic correction filter obtained by unwarping the warped filter; wherein an audio signal, filtered by the filtering means comprised of the room acoustic correction filter, is received substantially distortion-free at each of the expected listener positions. The weighted average response may be determined by a pattern recognition means (at least one of a hard c-means clustering system, a fuzzy c-means clustering system, or an adaptive learning system), and the warping is achieved by an all-pass filter. The warped filter includes an inverse of the lower order spectral model (such as a frequency pole-zero model or an LPC model). Thus, filtering each of the acoustical responses with the room acoustical correction filter provides a substantially flat magnitude response at each of the listener positions.

[0037] In another embodiment of the present invention, a method for correcting room acoustics at multiple-listener positions comprises: (i) warping each room acoustical response, said each room acoustical response obtained at each expected listener position; (ii) clustering each of the warped room acoustical response into at least one

cluster, wherein each cluster includes a centroid; (iii) forming a general response from the at least one centroid; (iv) inverting the general response to obtain an inverse response; (v) obtaining a lower order spectral model of the inverse response; (vi) unwarping the lower order spectral model of the inverse response to form the room acoustic correction filter; wherein the room acoustic correction filter corrects the room acoustics at the multiple-listener positions.

[0038] BRIEF DESCRIPTION OF THE DRAWINGS

[0039] FIG. **1** shows the basics of sound propagation characteristics from a loudspeaker to a listener in an environment such as a room, movie-theater, home-theater, automobile interior;

[0040] FIG. **2** shows an exemplary depiction of two responses measured in the same room a few feet apart;

[0041] FIG. **3** shows frequency response plots that justify the need for performing multiple-listener equalization;

[0042] FIG. **4** depicts a block diagram overview of a multiple-listener equalization system (i.e., the room acoustical correction system), including the room acoustical correction filter and the room acoustical responses at each expected listener position;

[0043] FIG. **5** shows the motivation for using the weighted averaging process (or means) for performing multiple-listener equalization;

[0044] FIG. **6** shows one embodiment for designing the room acoustical correction filter;

[0045] FIG. **7** shows the original frequency response plots obtained at six listener positions (with one loudspeaker);

[0046] FIG. **8** shows the corrected (equalized) frequency response plots on using the room acoustical correction filter according to one aspect of the present invention;

[0047] FIG. **9** is a flow chart to determine the room acoustical correction filter according to one aspect of the invention;

[0048] FIG. **10** is a flow chart to determine the room acoustical correction filter according to another aspect of the invention;

[0049] FIG. **11** is a flow chart to determine the room acoustical correction filter according to another aspect of the invention;

[0050] FIG. **12** is a flow chart to determine the room acoustical correction filter according to another aspect of the invention;

[0051] FIG. **13** is a pole zero plot of a signal to be modeled using Linear Predictive Coding (LPC);

[0052] FIG. **14** is a plot depicting the frequency response of the signal of FIG. 13 along with the approximation of the response with various order of the LPC algorithm;

[0053] FIG. **15** shows the implementation for warping a room acoustical response;

[0054] FIG. **16** is a figure showing different curves associated with different warping parameters for frequency axis warping;

[0055] FIG. **17** is a figure showing different frequency resolutions achieved for different warping parameters;

[0056] FIG. **18** is an example of a magnitude response of an acoustical impulse response;

[0057] FIG. **19** is the warped magnitude response corresponding to the magnitude response in FIG. **18**;

[0058] FIG. **20** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to one aspect of the present invention;

[0059] FIG. **21** are exemplary frequency response plots obtained at six listener positions;

[0060] FIG. **22** show the frequency response plots at the six listener positions of FIG. **21** that were corrected by using 512 tap room acoustical correction filter according to one aspect of the present invention;

[0061] FIG. **23** are exemplary frequency response plots obtained at six listener positions; and

[0062] FIG. **24** show the frequency response plots at the six listener positions of FIG. **23** that were corrected by using 512 tap room acoustical correction filter according to one aspect of the present invention.

[0063] FIG. **25** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to another aspect of the present invention.

[0064] DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0065] FIG. **1** shows the basics of sound propagation characteristics from a loudspeaker (shown as only one for ease in depiction) **20** to multiple listeners (shown to be six in an exemplary depiction) **22** in an environment **10**. The direct path of the sound, which may be different for different listeners, is depicted as **24**, **25**, **26**, **27**, **28**, **29**, and **30** for listeners one through six. The reflected path of the sound, which again may be different for different listeners, is depicted as **31** and is shown only for one listener here (for ease in depiction).

[0066] The sound propagation characteristics may be described by the room acoustical impulse response, which is a compact representation of how sound propagates in an environment (or enclosure). Thus, the room acoustical response includes the direct path and the reflection path components of the sound field. The room acoustical response may be measured by a microphone at an expected listener position. This is done by, (i) transmitting a stimulus signal (e.g., a logarithm chirp, a broadband noise signal, a maximum length signal, or any other signal that sufficiently excites the enclosure modes) from the loudspeaker, (ii) recording the signal received at an expected listener position, and (iii) removing (deconvolving) the response of the microphone (also possibly removing the response associated with the loudspeaker).

[0067] Even though the direct and reflection path taken by the sound from each loudspeaker to each listener may appear to be different (i.e., the room acoustical

impulse responses may be different), there may be inherent similarities in the measured room responses. In one embodiment of the present invention, these similarities in the room responses, between loudspeakers and listeners, may be used to form a room acoustical correction filter.

[0068] FIG. **2** shows an exemplary depiction of two responses measured in the same room a few feet apart. The left panels **60** and **64** show the time domain plots, whereas the right panels **68** and **72** show the magnitude response plots. The room acoustical responses were obtained at two expected listener positions, in the same room. The time domain plots, **60** and **64**, clearly show the initial peak and the early/late reflections. Furthermore, the time delay associated with the direct path and the early and late reflection components between the two responses exhibit different characteristics.

[0069] Furthermore, the right panels, **68** and **72**, clearly show a significant amount of distortion introduced at various frequencies. Specifically, certain frequencies are boosted (e.g., 150 Hz in the bottom right panel **72**), whereas other frequencies are attenuated (e.g., 150 Hz in the top right panel **68**) by more than 10 dB. One of the objectives of the room acoustical correction filter is to reduce the deviation in the magnitude response, at all expected listener positions simultaneously, and make the spectrum envelopes flat. Another objective is to remove the effects of early and late reflections, so that the effective response (after applying the room acoustical correction filter) is a delayed Kronecker delta function, $\delta(n)$, at all listener positions.

[0070] FIG. **3** shows frequency response plots that justify the need for performing multiple-listener room acoustical correction. Shown therein is the fact that, if an inverse

filter is designed that "flattens" the magnitude response, at one position, then the response is degraded significantly in the other listener position.

[0071] Specifically, the top left panel **80** in FIG. **3** is the correction filter obtained by inverting the magnitude response of one position (i.e., the response of the top right panel **68**) of FIG. **2**. Upon using this filter, clearly the resulting response at one expected listener position is flattened (shown in top right panel **88**). However, upon filtering the room acoustical response of the bottom left panel **84** (i.e., the response at another expected listener position) with the inverse filter of panel **80**, it can be seen that the resulting response (depicted in panel **90**) is degraded significantly. In fact there is an extra 10 dB boost at 150 Hz. Clearly, a room acoustical correction filter has to minimize the spectral deviation at all expected listener positions simultaneously.

[0072] FIG. **4** depicts a block diagram overview of the multiple-listener equalization system. The system includes the room acoustical correction filter **100**, of the present invention, which preprocesses or filters the audio signal before transmitting the processed (i.e., filtered) audio signal by loudspeakers (not shown). The loudspeakers and room transmission characteristics (simultaneously called the room acoustical response) are depicted as a single block **102** (for simplicity). As described earlier, and is well known in the art, the room acoustical responses are different for each expected listener position in the room.

[0073] Since the room acoustical responses are substantially different for different source-listener positions, it seems natural that whatever similarities reside in the responses be maximally utilized for designing the room acoustical correction filter **100**.

Accordingly, in one aspect of the present invention, the room acoustical correction filter **100** may be designed using a "similarity" search algorithm or a pattern recognition algorithm/system. In another aspect of the present invention, the room acoustical correction filter **100** may be designed using a weighted average scheme that employs the similarity search algorithm. The weighted average scheme could be a recursive least squares scheme, a scheme based on neural-nets, an adaptive learning scheme, a pattern recognition scheme, or any combination thereof.

[0074] In one aspect of the present invention, the "similarity" search algorithm is a c-means algorithm (e.g., the hard c-means of fuzzy c-means, also called k-means in some literatures). The motivation for using a clustering algorithm, such as the fuzzy c-means algorithm, is described with the aid of FIG. **5**.

[0075] FIG. **5** shows the motivation for using the fuzzy c-means algorithm for designing the room acoustical correction filter **100** for performing simultaneous multiple-listener equalization. Specifically, there is a high likelihood that the direct path component of the room acoustical response associated with listener 3 is similar (in the Euclidean sense) to the direct path component of the room acoustical response associated with listener 1 (since listener 1 and 3 are at same radial distance from the loudspeaker). Furthermore, it may so happen that the reflective component of listener 3 room acoustical response may be similar to the reflective component of listener 2 room acoustical response (due to the proximity of the listeners). Thus, it is clear that if responses 1 and 2 are clustered separately, due to their "dissimilarity", then response 3 should belong to the both clusters to some degree. Thus, this clustering approach permits an intuitively "sound" model for performing room acoustical correction.

[0076] The fuzzy c-means clustering procedures use an objective function, such as a sum of squared distances from the cluster room response prototypes, and seek a grouping (cluster formation) that extremizes the objective function. Specifically, the objective function, $J_\kappa(.,.)$, to minimize in the fuzzy c-means algorithm is:

[0077]

$$
\begin{aligned}
J_\kappa(U_{c\times N}, \underline{\hat{h}}^*) &= \sum_{i=1}^{c}\sum_{k=1}^{N}(\mu_i(\underline{h}_k))^\kappa (d_{ik})^2 \\
\mu_i(\underline{h}_k) &\in U_{c\times N}; \quad \mu_i(\underline{h}_k) \in [0,1] \\
\underline{\hat{h}}^* = (\underline{\hat{h}}_1^*, \underline{\hat{h}}_2^*, \ldots, \underline{\hat{h}}_c^*); \quad d_{ik}^2 &= \|\underline{h}_k - \underline{\hat{h}}_i^*\|^2
\end{aligned}
$$

[0078] In the above equation, $\underline{\hat{h}_i}^*$, denotes the i-th cluster room response prototype (or centroid), $\underline{h}_k$ is the room response expressed in vector form (i.e., $\underline{h}_k = (h_i(n); n = 0,1,\ldots) = (h_i(0), h_i(1), \ldots, h_i(M-1))^T$ and T represents the transpose operator), *N* is the number of listeners, *c* denotes the number of clusters (*c* was selected as $\sqrt{N}$, but could be some value less than *N*), $\mu_i(\underline{h}_k)$ is the degree of membership of acoustical response *k* in cluster *i*, d_{ik} is the distance between centroid $\underline{\hat{h}_i}^*$ and response $\underline{h}_k$, and κ is a weighting parameter that controls the fuzziness in the clustering procedure. When $\kappa = 1$, fuzzy c-means algorithm approaches the hard c-means algorithm. The parameter

κ was set at 2 (although this could be set to a different value between 1.25 and infinity). It can be shown that on setting the following:

[0079] $$\partial J_2(.)/\partial \underline{\hat{h}}_i^* = 0 \text{ and } \partial J_2(.)/\partial \mu_i(\underline{h}_k) = 0$$

[0080] yields:

[0081]
$$\begin{aligned} \underline{\hat{h}}_i^* &= \frac{\sum_{k=1}^{N} (\mu_i(\underline{h}_k))^2 \underline{h}_k}{\sum_{k=1}^{N} (\mu_i(\underline{h}_k))^2} \\ \mu_i(\underline{h}_k) &= [\sum_{j=1}^{c} (\frac{d_{ik}^2}{d_{jk}^2})]^{-1} = \frac{\frac{1}{d_{ik}^2}}{\sum_{j=1}^{c} \frac{1}{d_{jk}^2}}; \quad i = 1, 2, ..., c; k = 1, 2, ..., N \end{aligned}$$

[0082] An iterative optimization was used for determining the quantites in the above equations. In the trivial case when all the room responses belong to a single cluster, the single cluster room response prototype $\underline{\hat{h}}_i^*$ is the uniform weighted average (i.e., a spatial average) of the room responses since, $\mu_i(\underline{h}_k)$=1, for all k. In one aspect of the present invention for designing the room acoustical correction filter, the resulting room response formed from spatially averaging the individual room responses at multiple locations is stably inverted to form a multiple-listener room acoustical correction filter. In reality, the advantage of the present invention resides in applying non-uniform weights to the room acoustical responses in an intelligent manner (rather than applying equal weighting to each of these responses).

[0083] After the centroids are determined, it is required to form the room acoustical correction filter. The present invention includes different embodiments for designing multiple-listener room acoustical correction filters.

[0084] A. Spatial Equalizing Filter Bank:

[0085] FIG. **6** shows one embodiment for designing the room acoustical correction filter with a spatial filter bank. The room responses, at locations where the responses need to be corrected (equalized), may be obtained a priori. The c-means clustering algorithm is applied to the acoustical room responses to form the cluster prototypes. As depicted by the system in Fig. **6**, based on the location of a listener "i", an algorithm determines, through the imaging system, to which cluster the response for listener "i" may belong. In one aspect of the invention, the minimum phase inverse of the corresponding cluster centroid is applied to the audio signal, before transmitting through the loudspeaker, thereby correcting the room acoustical characteristics at listener "i".

[0086] B. Combining the Acoustical Room Responses using Fuzzy Membership Functions:

[0087] The objective may be to design a single equalizing or room acoustical correction filter (either for each loudspeaker and multiple-listener set, or for all loudspeakers and all listeners), using the prototypes or centroids $\underline{\hat{h}}_i^*$. In one embodiment of the present invention, the following model is used:

[0088]
$$\underline{h}_{final} = \frac{\sum_{j=1}^{c}(\sum_{k=1}^{N}(\mu_j(\underline{h}_k))^2)\underline{\hat{h}}_j^*}{\sum_{j=1}^{c}(\sum_{k=1}^{N}(\mu_j(h_k))^2)}$$

[0089] $\underline{h}_{final}$ is the general response (or final prototype) obtained by performing a weighted average of the centroids $\underline{\hat{h}}_i^*$. The weights for each of the centroids, $\underline{\hat{h}}_i^*$, is determined by the "weight" of that cluster "i", and is expressed as:

[0090] $$weight_i = \frac{\sum_{k=1}^{N} \mu_i (\underline{h}_k)^2}{\sum_{i=1}^{c} \sum_{k=1}^{N} \mu_i (\underline{h}_k)^2}$$

[0091] It is well known in the art that any signal can be decomposed into its minimum-phase part and its all-pass part. Thus,

[0092] $$h_{final}(n) = h_{\min, final}(n) \otimes h_{ap,final}(n)$$

[0093] The multiple-listener room acoustical correction filter is obtained by either of the following means, (i) inverting $\underline{h}_{final}$, (ii) inverting the minimum phase part, $\underline{h}_{\min, final}$, of $\underline{h}_{final}$, (iii) forming a matched filter $\underline{h}_{ap,final}^{matched}$ from the all pass component (signal), $\underline{h}_{ap,final}$, of $\underline{h}_{final}$, and filtering this matched filter with the inverse of the minimum phase signal $\underline{h}_{\min, final}$. The matched filter may be determined, from the all-pass signal as follows:

[0094] $$h_{ap,final}^{matched}(n) = h_{ap,final}(-n + \Delta)$$

[0095] Δ is a delay term and it may be greater than zero. In essence, the matched filter is formed by time-domain reversal and delay of the all-pass signal.

[0096] The matched filter for multiple-listener environment can be designed in several different ways: (i) form the matched filter for one listener and use this filter for all

listeners, (ii) use an adaptive learning algorithm (e.g., recursive least squares, an LMS algorithm, neural networks based algorithm, etc.) to find a "global" matched filter that best fits the matched filters for all listeners, (iii) use an adaptive learning algorithm to find a "global" all-pass signal, the resulting global signal may be time-domain reversed and delayed to get a matched filter.

[0097] FIG. **7** shows the frequency response plots obtained on using the room acoustical correction filter for one loudspeaker and six listener positions according to one aspect of the present invention. Only one set of loudspeaker to multiple-listener acoustical responses are shown for simplicity. Large spectral deviations and significant variation in the envelope structure can be seen clearly due to the differences in acoustical characteristics at the different listener positions.

[0098] FIG. **8** shows the corrected (equalized) frequency response plots on using the room acoustical correction filter according to one aspect of the present invention (viz., inverting the minimum phase part, $\underline{h}_{\min,final}$, of $\underline{h}_{final}$, to form the correction filter). Clearly, the spectral deviations have been substantially minimized at all of the six listener positions, and the envelope is substantially uniform or flattened thereby substantially eliminating or reducing the distortions of a loudspeaker transmitted audio signal. This is because the multiple-listener room acoustical correction filter compensates for the poor acoustics at all listener positions simultaneously.

[0099] FIGs. **9-12** are the flow charts for four exemplary depictions of the invention.

[00100] In another embodiment of the present invention, the pattern recognition technique can be used to cluster the direct path responses separately, and the reflective

path components separately. The direct path centroids can be combined to form a general direct path response, and the reflective path centroids may be combined to form the general reflective path response. The direct path general response and the reflective path general response may be combined through a weighted process. The result can be used to determine the multiple-listener room acoustical correction filter (either by inverting the result, or the stable component, or via matched filtering of the stable component).

[00101] The filter in the above case was an 8192 finite impulse response (FIR) filter. This filter was obtained from 8192-coefficient impulse responses sampled at 48kHz sampling frequency. In order for realizable filters that can be implemented in a cost effective manner for real-time DSP applications (e.g., home-theater, automobiles, etc.), the number of filter coefficients should be substantially reduced without substantial changes in the results (subjective and objective).

[00102] Accordingly, in one embodiment of the present invention, a lower order multiple location (listener) equalization filter is designed by (i) warping the room responses to the Bark scale using the concepts from, (ii) performing data clustering to determine similarities between room responses (essentially a non-uniform weighting approach) for finding a "prototype" response, (iii) fitting a lower order spectral model (e.g., a pole zero model or an LPC model), (iv) inverting the LPC model to determine a filter in the warped domain, and (v) unwarping the filter onto the linear axis to get the equalizing filter. FIG. **20** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to this aspect of the present invention.

[00103] Accordingly, in another embodiment of the present invention, a lower order multiple location (listener) equalization filter is designed by (i) warping the room responses to the Bark scale using the concepts from, (ii) performing data clustering to determine similarities between room responses (essentially a non-uniform weighting approach) for finding a "prototype" response, (iii) inverting the prototype response as found y the non-uniform weighting approach of the clustering algorithm, (iv) fitting a lower order spectral model (e.g., a pole zero model or an LPC model) to the prototype (or general) response to form a filter in the warped domain,and (iv) unwarping the filter onto the linear axis to get the equalizing filter. FIG. **25** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to this aspect of the present invention.

[00104] <u>Spectral Modelling with LPC:</u>

[00105] Linear predictive coding is used widely for modelling speech spectra with a fairly small number of parameters called the predictor coefficients. It can also be applied to model room responses in order to develop low order equalization filters. As shown through the following example, effective low order inverse filters can be formed through LPC modelling.

[00106] The error equation *e(n)*, for a signal *s(n)* (to be modeled by $\tilde{s}(n)$), governing the all-pole LPC model of order *p* and predictor coefficients a_k is expressed as:

$$e(n) = s(n) - \tilde{s(n)} = s(n) - \sum_{k=1}^{p} a_k s(n-k)$$

[00107]

[00108] Specifically, Fig. **13** shows a stable minimum phase signal having five zeros and four poles, whereas FIG. **14** is a plot depicting the frequency response of the signal of FIG. **13** along with the approximation of the response with various orders (i.e., number of predictor coefficients being 16, 32, and 128) of the LPC algorithm.

[00109] The LPC transfer function $H_1(z)$, which employs an all-pole model, that approximates the signal, $s(n)$, transform $S(z)$ is expressed as:

[00110] $$H_1(z) = \frac{K}{\sum_{k=1}^{p} a_k z^{-k}}$$

[00111] where K is an appropriate gain term. Alternative models (such as pole-zero models) can be used, and these are expressed as:

[00112] $$H_2(z) = \frac{\sum_{l=1}^{r} b_k z^{-k}}{\sum_{k=1}^{p} a_k z^{-k}}$$

[00113] In addition, the all-pole (LPC) model $H_1(z)$ and/or the pole-zero model $H_2(z)$ can be frequency weighted to approximate the signal transform $S(z)$ selectively in specific frequency regions using the following objective function that is to be minimized with respect to θ and the frequency weighting term $W(e^{j\omega}$:

[00114] $J(\theta) = \left\| A(e^{j\omega})S(e^{j\omega}) - B(e^{j\omega}) \right\|_2^2 W(e^{j\omega})$

[00115] where:

[00116] $A(z) = \dfrac{K_1}{\sum_{k=1}^{p} a_k z^{-k}}; \quad B(z) = \sum_{l=1}^{r} b_k z^{-k}; \quad \theta = [a_1,\ldots,a_p,b_1,\ldots,b_r]$

[00117] FIG. **15** shows the implementation for warping, through the bilinear conformal map, a room acoustical response using an all-pass filter chain. The basic idea for warping is done using an FIR chain having all-pass blocks (with all-pass or warping coefficients λ), instead of conventional delay elements. When an all-pass filter, $D_1(z)$, is used, the frequency axis is warped and the resulting frequency response is obtained at non-uniformly sampled points along the unit circle. Thus, for warping

$$D_1(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}}$$

[00118]

[00119] The group delay of $D_1(z)$ is frequency dependent, so that positive values of the warping coefficient λ yield higher frequency resolutions in the original response for low frequencies, whereas negative values of λ yield higher resolutions in the frequency response at high frequencies.

[00120] Clearly, the cascade chain of all-pass filters result in an infinite duration sequence. Typically a windowing is employed that truncates this infinite duration sequence to a finite duration to yield an approximation.

[00121] Warping via a bilinear conformal map and based on the all-pass transformation to the psycho-acoustic Bark frequency scale can be obtained by the following relation between the warping parameter λ and the sampling frequency f_s:

[00122]
$$\lambda = 0.8517[arctan(0.06583 f_s)]^{1/2} - 0.1916$$

[00123] FIG. **16** is a figure showing different curves associated with different warping parameters, λ, for transformation of the frequency response via frequency warping. Positive values of the warping parameter map low frequencies to high frequencies (which translates into stretching the frequency response), where negative values of the warping parameter map high frequencies to low frequencies. During the unwarping stage the warping parameter is selected to be $-\lambda$.

[00124] FIG. **17** is a figure showing different frequency resolutions for positive warping parameters.

[00125] FIG. **18** is an example of a magnitude response of an acoustical impulse response, whereas FIG. **19** is the warped magnitude response corresponding to the magnitude response in FIG. **18** (with λ=0.78).

[00126] FIG. **20** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to one aspect of the present invention, showing several steps. The first step involves measuring the room impulse response at each of the expected listener positions. Subsequently, the room responses are warped based on the warping parameter λ before lower order spectral fitting. Warping is important

since it is important to get a good resolution, particularly at lower frequencies, so that the lower order LPC spectral model, used in the subsequent stage, can achieve a good fit to a frequency response in the lower frequencies (below 6 kHz). After warping each response, weighting, using some non-uniform weighting method or by a pattern recognition method or fuzzy clustering method or through a simple energy averaging (i.e., root-mean-square RMS averaging) method, is done to the warped responses to obtain a general response or a prototype response (e.g., as in paragraph [0080] where $\underline{h}_k$ are the warped responses and the general response in the warped domain is $\hat{\underline{h}}_i^*$). After determining the general response, a lower order model (e.g., the LPC model, a pole-zero model, a frequency weighted LPC or pole-zero model) may be used to model the general response with a small number of coefficients (e.g., the predictor coefficients a_k). The resulting impulse response from the LPC model is then inverted to get a filter in the warped domain. An unwarping stage, with warping parameter $-\lambda$, unwarps the frequency response of the filter in the warped domain to give a room acoustical correction filter in the linear frequency domain. The first L taps of the room acoustical correction filter are selected (where L<P, P being the length of the room response). Thus, conventional Fast Fourier Transform algorithms may be used for real-time signal processing and filtering with the L taps of the room acoustical correction filter.

[00127] FIG. **21** are exemplary frequency response plots obtained at six listener positions in a room for one loudspeaker, whereas FIG. **22** shows the frequency response plots at the six listener positions of FIG. **21** that were corrected by using L=512 tap room acoustical correction filter (with k=512 predictor coefficients in the LPC)

according to one aspect of the present invention using $\lambda=0.78$. Each subplot, in each figure, corresponds to the frequency response at one listener position. Clearly, there is a significant amount of correction as the room correction filter minimizes the magnitudes of the peaks and dips that cause significant degradation in the perceived audio quality. The resulting frequency response at the six listener positions is substantially flat as can be seen through FIG. **22**.

[00128] FIG. **23** are exemplary frequency response plots for another system in a room obtained at six listener positions for another loudspeaker, whereas FIG. **24** show the frequency response plots at the six listener positions of FIG. **23** that were corrected by using L=512 tap room acoustical correction filter according to one aspect of the present invention.

[00129] FIG. **25** is a block diagram for achieving low filter orders for performing multiple-listener equalization according to another aspect of the present invention. In this embodiment, the inverse filter is first determined using at least the minimum phase part of the prototype response. A lower order spectral model (e.g., LPC) is then fitted to the inverse response to obtain a lower order warped filter. The warped filter is unwarped to get the room acoustical correction filter in the linear frequency domain. The first L taps of this filter may be selected for real-time room acoustical equalization.

[00130] The description of exemplary and anticipated embodiments of the invention has been presented for the purposes of illustration and description purposes. They are not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations are possible in light of the teachings herein. For example,

the number of loudspeakers and listeners may be arbitrary (in which case the correction filter may be determined (i) for each loudspeaker and multiple-listener responses, or (ii) for all loudspeakers and multiple-listener responses). Additional filtering may be done to shape the final response, at each listener, such that there is a gentle roll-off for specific frequency ranges (instead of having a substantially flat response).